

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2003, 10:33:25 ; Search time 2192 Seconds  
(without alignments)  
12279.607 Million cell updates/sec

Title: US-10-037-311A-2

Perfect score: 1662

Sequence: 1 atgagtcagattcgtacag.....ggggacttaagctagtatga 1662

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpi:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	481.4	29.0	483	9 AI994677	AI994677 701499231
2	440.8	26.5	559	14 R90192	R90192 16547 Lambd
3	331	19.9	416	10 AV805720	AV805720 AV805720
4	309.6	18.6	668	17 B77193	B77193 T31J1TR TAM
5	296	17.8	507	9 AA605483	AA605483 30519 Lam
6	281	16.9	645	10 BE319670	BE319670 NF017D12R

7	280	16.8	597	13	BI416307	BI416307 LJNEST4b4
8	272	16.4	443	10	AV807820	AV807820 AV807820
9	259.8	15.6	559	10	AW719438	AW719438 LJNEST4f1
10	259.6	15.6	671	17	B23728	B23728 F16M20TR IG
11	252.6	15.2	508	12	BF597909	BF597909 sv02f04.y
12	250.2	15.1	561	12	BF009322	BF009322 ss76g10.y
13	245.2	14.8	605	10	BE346170	BE346170 sp19b08.y
14	235.6	14.2	314	10	AV831440	AV831440 AV831440
15	232.4	14.0	786	17	BI8679	BI8679 F16A14-T7 I
16	224.4	13.5	535	12	BG356747	BG356747 OV2_9.E02
17	209.2	12.6	505	12	BG045011	BG045011 saa36d05.
18	205.4	12.4	552	13	BI969461	BI969461 GM830008A
19	199.2	12.0	658	10	AW693268	AW693268 NF064A03S
20	188.2	11.9	592	17	BH555955	BH555955 BOHFU09TR
21	188.8	11.4	570	13	BI425718	BI425718 sab71b08.
22	187	11.3	404	17	BH862014	BH862014 SALK_0885
23	187	11.3	594	12	BE802468	BE802468 sr31f08.y
24	181.2	10.9	470	13	BI969504	BI969504 GM830008A
25	178.2	10.7	597	13	BI946374	BI946374 O1117 lea
26	178	10.7	380	10	BE662954	BE662954 EST00099
27	172	10.3	418	14	BQ985962	BQ985962 QGE8D08.y
28	172	10.3	430	14	BQ980670	BQ980670 QGE11L07.
29	172	10.3	448	14	BQ990773	BQ990773 QGF21B03.
30	172	10.3	465	14	BQ873962	BQ873962 QG14A14.y
31	170.4	10.3	373	14	BQ984710	BQ984710 QGE4E05.y
32	168.4	10.1	562	13	BM525521	BM525521 sal28h09
33	161.8	9.7	613	10	BE361458	BE361458 DG1_72_E0
34	159.4	9.6	649	10	AV836113	AV836113 AV836113
35	158.6	9.5	992	10	BE040029	BE040029 OC10F08 O
36	151.6	9.1	323	17	AZ130387	AZ130387 OSJNB010
37	150.8	9.1	740	12	BE823653	BE823653 GM700021A
38	149.8	9.0	812	17	BH495933	BH495933 BOHNS54TF
39	142.4	8.6	301	17	AQ796429	AQ796429 nbxb0066A
40	142	8.5	415	10	AW704030	AW704030 SK27A04.y
41	141.4	8.5	529	12	BG357064	BG357064 OV2_9.E02
42	140.8	8.5	663	13	BJ314543	BJ314543 BJ314543
43	140.2	8.4	564	10	AW720399	AW720399 LJNEST22C
44	139.6	8.4	377	17	BH548166	BH548166 BQGW62TR
45	139.6	8.4	529	17	AQ250893	AQ250893 T26M12-17

#### ALIGNMENTS

RESULT 1	AI994677	483 bp	mRNA	linear	EST 08-SEP-1999
AI994677	701499231	A. thaliana, Ohio State clone set	Arabidopsis thaliana		
LOCUS	701499231	A. thaliana, Ohio State clone set	Arabidopsis thaliana		
DEFINITION	701499231	A. thaliana, Ohio State clone set	Arabidopsis thaliana		
ACCESSION	AI994677	483 bp	mRNA	linear	EST 08-SEP-1999
VERSION	AI994677	483 bp	mRNA	linear	EST 08-SEP-1999
KEYWORDS	EST.				
SOURCE	AI994677.1	GI:5841582			
ORGANISM	Arabidopsis thaliana				
REFERENCE	Chen, J., Momiya, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorsone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, N., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, F., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrta, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.				
AUTHORS	Arabidopsis thaliana				
TITLE	Unpublished (1999)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: David Smoller, Ph.D. Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc. 4633 World Parkway Circle, St. Louis, MO 63134, USA Tel: 877-577-2733 Fax: 314-427-3324 Email: service@genomesystems.com.				

**COMMENT**  
 Contact: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
 Lansing, MI  
 Tel: 517-353-0034  
 Fax: 517-353-9168  
 Email: 22313tcn@lhm.cl.msu.edu  
 Seq primer: T7 dye primer.  
**FEATURES**  
 source  
 Location/Qualifiers  
 1..559  
 /organism="Arabidopsis thaliana"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="191A67"  
 /clone\_lid="Lambda-PRL2"  
 /note="Vector: lambda Zip-Lox; Site\_1: Sal; Site\_2: Not;  
 Lambda PRL2 is a cDNA library derived from equal  
 quantities of 4 pools of mRNA. The mRNA sources were 1)  
 day germinated etiolated seedlings; 2) tissue culture  
 grown roots; 3) staged plants half with 24 hour light  
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques. The vector is pRL's lambda zip-Lox. The cDNA  
 inserts were directionally cloned with Sal-Not arms using  
 oligo dT primed cDNA."  
 BASE COUNT 151 a 107 c 130 g 158 t 13 others  
 ORIGIN  
 Query Match 26.5%; Score 440.8; DB 14; Length 559;  
 Best Local Similarity 95.1%; Pred. No. 1.1e-118;  
 Matches 451; Conservative 0; Mismatches 23; Indels 0; Gaps 0  
 QY 805 GGAGATCATGATAAGATGTTCTCTCTGAGAGAGACCAACATTCATCGGAAGTCCCT 864  
 Db 1 GGAGATCATGATAAGATGTTCTCTCTGAGAGAGACCAACATTCATCGGAAGTCCCT 60  
 QY 865 TGGTTGATGTTTAAACAGACAATTACTTTGTTCCATCTCTGTGGTTAATACCGGGTTTC 924  
 Db 61 TGGTTGATGTTTAAACAGACANTTACTTTGTTCCATCTCTGTGGTTAATACCGGGTTTC 120  
 QY 925 GATGATGAACATAACAAGCTATTCACCAGAAAGGACATGCTCTTTCATCATTAGTAGTAGG 984  
 Db 121 GATGATGAACATAACAAGCTATTCACCAGAAAGGACATGTTTTCATCATTAGTAGTAGG 180  
 QY 985 TATCTTTTTCACCAACATAACCAAGTATGGGCTTAGTCTACATAGATACACGAGCTTAC 1044  
 Db 181 TATCTTTTTCACCAACATAACCAAGTATGGGCTTAGTCTACATAGATACACGAGCTTAC 240  
 QY 1045 TTATCGCATGCGGATGAGAAGATTCGGGATTCAAGTAAGAGTTTTCGATGAAGACCCGGGT 1104  
 Db 241 TTATCGCATGCGGATGAGAAGATTCGGGATTCAAGTAAGAGTTTTCGATGAAGACCCGGGT 300  
 QY 1105 CCATTTCAGCATGTGATGGATCAGATTTTCATCTTGTACTCAAAAAGAGAAATCTCTACCT 1164  
 Db 301 CCATTTCAGCATGTGATGGATCAGATTTTCATCTTGTACTCAAAAAGAGAAATCTCTACCT 360  
 QY 1165 GAAGTAGACACATAGTACGAGATCTCCGATATTAATACCCCAACACAAAGCCGTCG 1224  
 Db 361 GAAGTAGACACATAGTACGAGATCTCCGATATTAATACCCCAACACAAAGCCGTCG 420  
 QY 1225 CTTGTACATCTTTGAACCGGGTTACCGGGAGAACTTAACAGATCATATTGG 1278  
 Db 421 CTTGTACATCTTTGAACCGGGTTACCGGGAGAACTTAACAGATCATATTGG 474  
**RESULT 3**  
 AV805720/c  
 LOCUS AV805720  
 DEFINITION RAFL9 Arabidopsis thaliana cDNA clone linear EST 29-MAR-2001  
 mRNA sequence.  
 ACCESSION AV805720  
 VERSION AV805720.1 GI:19830705

KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 416)  
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,  
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,  
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.  
and Shinozaki,K.  
Large scale analysis of Arabidopsis full-length cDNA (2002b)  
UNPUBLISHED (2002)  
CONTACT: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4339  
Fax: 81-298-36-9060  
Email: mseki@tc.riken.go.jp

TITLE An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et  
al., submitted for publication) digested with BamHI and SalI. This  
clone is in a modified pBluescript vector. Please visit our web  
site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further  
details.

FEATURES  
source Location/Qualifiers  
1..416  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/clone="RAFL09-43-P18"  
/clone\_lib="RAFL9"  
/dev\_stage="plants at various developmental stages from  
germination to mature seeds"  
/lab\_host="DH10B"  
/note="Site\_1: BamHI; Site\_2: SalI; subjected to  
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24  
hr) treatments"  
BASE COUNT 138 a 94 c 74 g 110 t  
ORIGIN

Query Match 19.9%; Score 331; DB 10; Length 416;  
Best Local Similarity 100.0%; Pred. No. 2.5e-86;  
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1332 AGGTTATCAGCAGACCGGAAAAAGATGCATTAATGGCAAGCTCTTGGGGAATGTATCT 1391  
Db 416 AGGTTATCAGCAGACCGGAAAAAGATGCATTAATGGCAAGCTCTTGGGGAATGTATCT 357

QY 1392 TTTCAGTTTGACAGATAATCTGTGACAAGTCTTGCTGCTACATTTGGATATGTAGCTCA 1451  
Db 356 TTTCAGTTTGACAGATAATCTGTGACAAGTCTTGCTGCTACATTTGGATATGTAGCTCA 297

QY 1452 AGGCTTTGAGAGGTTTAAAGCCTTTGGATACTATAGACCCGAAAAACCGTACAACTCCCGA 1511  
Db 296 AGGCTTTGAGAGGTTTAAAGCCTTTGGATACTATAGACCCGAAAAACCGTACAACTCCCGA 237

QY 1512 TCCTTCGTGTGGTGGGTATGTGATGGAGCCTTTGTTTCCACTGCCCTCCATTCTATGA 1571  
Db 236 TCCTTCGTGTGGTGGGTATGTGATGGAGCCTTTGTTTCCACTGCCCTCCATTCTATGA 177

QY 1572 TTGTAAGCGAAACGGGTATGTACACGGGACACTAGTCTCTCATGTGAGACATTTGGA 1631  
Db 176 TTGTAAGCGAAACGGGTATGTACACGGGACACTAGTCTCTCATGTGAGACATTTGGA 117

QY 1632 GGATATCATGCTGGGACTTAAAGCTAGTATGA 1662  
Db 116 GGATATCATGCTGGGACTTAAAGCTAGTATGA 86

RESULT 4  
B77193

LOCUS B77193 668 bp DNA linear GSS 16-JAN-1998  
DEFINITION T31J1TR TAMU Arabidopsis thaliana genomic clone T31J1, DNA  
sequence.  
ACCESSION B77193  
VERSION B77193.1 GI:2773832  
KEYWORDS GSS.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana

REFERENCE  
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 668)  
Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K.,  
Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter  
J.C.  
A BAC End Sequence Database for Identifying Minimal Overlaps in  
Arabidopsis Genomic Sequencing. Update 3  
UNPUBLISHED (1997)  
Other\_GSSs: T31J1TF  
CONTACT: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: rounsley@tigr.org  
Seq primer: M13 Reverse  
Class: BAC ends  
High quality sequence stop: 668.

FEATURES  
source Location/Qualifiers  
1..668  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="T31J1"  
/clone\_lib="TAMU"  
/sex="hermaphrodite"  
/note="Vector: BelOBACII; Site\_1: HindIII; Site\_2: HindIII  
; Produced by Rod Wing"  
BASE COUNT 182 a 153 c 123 g 210 t  
ORIGIN

Query Match 18.6%; Score 309.6; DB 17; Length 668;  
Best Local Similarity 71.0%; Pred. No. 6e-80;  
Matches 430; Conservative 0; Mismatches 164; Indels 12; Gaps 1;

QY 486 TGATGTGAATGCAATATATGTTCTGTGGATTCTTTTAGCGGCTTAGGGAACAGGATACT 545  
Db 63 TGTGTGTAATGCAGATACATCTGTGGATTGCTTTATGGGCTAGGAAACAGATACT 122

QY 546 TTCTCTAGCCTCGGTTTCTTTACGGCTTTTACGGATAGAGTCTTGTGTTGACCG 605  
Db 123 TACTCTGTCTCTCTGTTTCTCTATGCTCTCTTACTGACGAGATCATGTTGTTGACCA 182

QY 606 AGGGAAGACATGGATGATCTCTTTTTCGAGCGCTTTTACGGATAGTGTGTTGCTTACC 665  
Db 183 ACGTACGGACATAAGTACCTCTCTGTGAGCCTTTTCCAGGTACTCTCTGGCTACTCCC 242

QY 666 TTTAGATTTCCCTTATGACTGATCAGTTTGTATGATTAATCAAGAATCATCTCGTTGTTA 725  
Db 243 TCTGGATTTTCCACTAACAGATCAATTAGATAGCTTCAACAAGGAATCTCCGGCTGTTA 302

QY 726 TGGATATATGGTGAAAGAAATCAGGTGATTGATCTAGGGGAAC-----TTTGTCT 773  
Db 303 CGGAACAATGTTGAAGAATCATGCCATTAATCAACTACACACAGAAAGCATCATCCCTC 362

QY 774 TCATCTTATCTCATCTTCTGTTTTCATGATTATGGAGATGATGTTCTTCTCTGA 833  
Db 363 GTACCTCTGCTTATCTTATTCACGATACGAGGATATGATAAGATGTTCTTCTGTA 422

QY 834 AGGAGACCAAAACATTCATCGGGAAGTCCCTGTGTTGATTGTTAAAAACAGACAATTAATT 893  
Db 423 AAGTGACCAAAATTCATCAGGCAAGTCCCTGTGTTGTTCTTCAACTCGAATCTTTACTT 482





ROSIDAE; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 443)

**REFERENCE**  
**\*AUTHORS** Seki,M., Narusawa,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Ono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shingawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.

**TITLE** Large scale analysis of Arabidopsis full-length cDNA (2002b)

**JOURNAL** Unpublished (2002)

**COMMENT** Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1, Koiyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: msekic@rtc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda E10-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit Our Web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

FEATURES	source	Location/Qualifiers
	1. .443	
		/organism="Arabidopsis thaliana"
		/db_xref="taxon:3702"
		/clone="RAFL09-52-C06"
		/Clone_lib="RAFL9"
		/dev_stage="plants at various developmental stages from germination to mature seeds"
		/lab_host="DH10B"
		/note="Site_1: BamHI; Site_2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
BASE COUNT	159 a 90 c 83 g 111 t	

Query Match	16.4%	Score 272;	DB 10;	Length 443;
Best Local Similarity	96.3%;	Pred. No. 6.4e-69;		
Matches 289;	Conservative 0;	Mismatches 10;	Indels 1;	Gaps 1;
QY	1364	ATGCGAAAGCCTCTGCGGAAATGATATCTTTTTCAGTTTTCACAGATAAATCTTTCACAACTG	1423	
Db	443	ATGCCAAGCTCTTTCGCGGAAATCATCTCTTTTGAGTTTGACAGATAATCTTGTCACAAGTG	384	
QY	1424	CTTGCTTCACATTTGGATATCTAGCTCAAGCTCTGGAGCTTTAAAGCCTTGGATACTCT	1483	
Db	383	CTTGCTTCACATTTGCAATTTATTTAGCTCAAGCTCTGGAGCTTTTAAACCTTTGGATACTCT	324	
QY	1484	ATAGACCCGAAACCGGTACAACTCCCGATCCTTCGTGTGGTTCGGGCTATGTCGATGGAGC	1543	
Db	323	ATAGCCGGAACCGGTACAACTCCGATCCTTCGTGTGGTTCGGGCTATGTCGATGGAGC	264	
QY	1544	CTTGTTTCCACTCGCCCTCCATTCATATCTT - AAAGCGAAACCGGTTATTCACACGGGA	1602	
Db	263	CTTGTTTCCACTCGCCCTCCATTCATGATTGTGAAGGCGAAACCGGTTATTCACACGGGA	204	
QY	1603	ACACTAGTTCCTCATGTGAGACATTTGTGAGATATACGCTGGGACCTTAAGCTAGTATGA	1662	
Db	203	ACACTAGTTCCTCATGTGAGACATTTGTGAGATATACGCTGGGACCTTAAGCTAGTATGA	144	

RESULT	9
LOCUS	AW719438
DEFINITION	LJNEST4f10r Lotus japonicus nodule library, mature and immature nodules Lotus japonicus cDNA 5', mRNA sequence.
ACCESSION	AW719438
VERSION	AW719438.1 GI:7613946
KEYWORDS	EST.
SOURCE	Lotus japonicus.
ORGANISM	Lotus japonicus Eukaryota; Viridiplantae; Streptophyta; Embryophyta Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.	
1 (bases 1 to 559)	
Freund, S., Stougaard, J. and Udvardi, M.	
Lotus japonicus root nodule ESTs: a tool for functional genomics	
Unpublished (2000)	
Contact: Udvardi MK	
Max Planck Institute of Molecular Plant Physiology	
Am Muehlenberg 1, 14476 Golm, Germany	
Fax: 49 331 567 8250	
Email: udvardi@mpimp-golm.mpg.de	
Seq primer: r7	
High quality sequence stop: 559.	

FEATURES	source
1. .559	
/organism="Lotus japonicus"	
/cultivar="Gifu (B-129)"	
/db_xref="taxon:34305"	
/clone_lib="Lotus japonicus nodule library, mature and immature nodules"	
/note="Organ: Nodule; Vector: pbluescript sk-; Site_1: ECORI; Site_2: XhoI; The library was prepared using mRNA extracted from a mixture of mature (pink) and immature (white) nodules of Lotus japonicus ecotype 'Gifu'. Nodules were induced by, and contained Mesorhizobium loti strain NZP2235."	
BASE COUNT	172 a 105 c 121 g 161 t
ORIGIN	

	Query Match.	15.6%	Score 259.8	DB 10	Length 559
	Best Local Similarity 67.0%	Pred. No. 2.8e-65			
	Matches 369	Conservative 0	Mismatches 182	Indels 0	Gaps 0
Qy	856	AAAGTCCCTTGGTTGATGTTTAAACACAGACAATACATTGTTTCCATCTCTGTGGTTAATA	915		
Db					
Db	5	AAAGTACCTTGGTTAGTGATGATGAACACAGATAAATATTTCTGTTCCATCTCTATTCTTGATG	64		
Qy	916	CCGGGTTTCCGATGATGATTAACACAGCTATTTCCACAGAAAGCGACTGCTTTTCATCAC	975		
Db					
Db	65	CCATCTTATGACGAGGAAGCTGAATGATCTCTTCCCAAGAAAGCAAGCTGTTCCATTTC	124		
Qy	976	TTAGTAGGTATCTTTTTCACCCAACATAACCAAGTATGCGGCTTTTACACACATAGATCTAC	1035		
Db					
Db	125	TTGGGTAGATATATGCTCCACCCACAAACAATGTGTGGGACTTGTGTGCAGATACTAT	184		
Qy	1036	GAAGCTTACTTATCGCATGCGGATGAGAAGATTGGGATTTCAAGTAAGAGTTTTCCATGAA	1095		
Db					
Db	185	CAAGCTTATTTAGCTAAAGTTGATCAAGAGTAGGCGATACAAATAGAGTGTGTGTGTTCCA	244		
Qy	1096	CACCGGGTCCATTTTCAGCATGTGATGGATCAGAGATTTCATCTCTACTCAAAAAGAGAAA	1155		
Db					
Db	245	GAACCTGGTCCATTTCACATGTTCTGGATCAGATCTTAGCTGTGACTTTGGAAGGAGAT	304		
Qy	1156	CTTCTACCTGAAGTACACACTAGTGGAGAGATCTCCGCAATGTTAATACCCCAACAC	1215		
Db					
Db	305	CTTTTATCCCATGTTGACCGTGAGCAGAATAATATTAGTTTCATCTGGACTACCAAGTCA	364		
Qy	1216	AAAGCGTGTGTGTGCACATCTTTCAACGCGGTTTACGGGAGAACTTTAAAGAGTATCTAT	1275		
Db					
Db	365	AAAGCTGTGTCATGATCATCTCTAAGCTGTGGTATTTTGAAGAGTAAGAGACATGTAT	424		
Qy	1276	TGGGAATATCCGACATCAACTGGAGAAATCATCGGTGTTCATACGCGAGGCCAAGAAGT	1335		
Db					
Db	425	TGGGAGTATCTTCCGTGACAGGAGAAAGTGTCTGGCGTTTTTCCAGCCCAAGCCATGAAGAG	484		
Qy	1336	TATCAGAGACCGAANAANAACATGCAATAATGCGAAAGCTCTTCCGGAANAATCTATCTTTTG	1395		
Db					
Db	485	TATCAGCAAGCGCAAGCAGATGCAACCTTAAAGCTTGGGCGAGAARTGATCTCTTAA	544		
Qy	1396	AGTTTGCACGA	1406		

```

Db 545 AGCTTACACA 555

RESULT 10
B23728
LOCUS
DEFINITION
  F16M20TR IGF Arabidopsis thaliana genomic clone F16M20, DNA
  sequence.
ACCESSION
  B23728
VERSION
  B23728.1
KEYWORDS
  GI:2509359
SOURCE
  thale cress.
ORGANISM
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
  1 (bases 1 to 671)
REFERENCE
  Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Coffey,E., Golden,K.,
  Johnson,K., Adams,M.D. and Venter,J.C.
  A BAC End Sequence Database for Identifying Minimal Overlaps in
  Arabidopsis Genomic Sequencing
  Unpublished (1997)
  Other_GSSs: F16M20TFB
  Contact: Steve Rounsley
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: rounsley@tigr.org
  Seq primer: M13 Reverse
  Class: BAC ends
  High quality sequence stop: 671.
  Location/Qualifiers
    1..671
      /organism="Arabidopsis thaliana"
      /strain="Columbia"
      /db_xref="taxon:3702"
      /clone="F16M20"
      /clone_lib="IGF"
      /note="hermaphrodite"
      /sex="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
        Produced by Thomas Altmann"
BASE COUNT 184 a 144 c 139 g 204 t
ORIGIN

Query Match 15.6%; Score 259.6; DB 17; Length 671;
Best Local Similarity 66.7%; Pred. No. 3.4e-65;
Matches 408; Conservative 0; Mismatches 189; Indels 15; Gaps 2;

QY 268 GATAAGCTTCTCGAGGGCTACTTGGTCTTCTGTTTTCATGAAGATTCTTGCCTTAGTAGG 327
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 GATAGACTAATAGGAGGGCTTTTAACCGCAGATTTCGATGAAGTTCTTGGTTGAGTAGG 119
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 328 TACC---AATCAGTTCATACCGTAAACCTTCACCTTACAGCCATCTCTTATCTCATC 384
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 TATCATAAACTTCTTGATCGAAGCCCTTACCATACAGCCGCTGATATCTTGTC 179
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 385 TCTAAGCTTAGAACTACGAAAAGCTTTCACAAGCGATGTGTCGGGTACTGAATCTTAC 444
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 TCGAAGCTTAGAAGCTATGATAGTCTTCACAAAGTTTGGCGTCCAGGACAAAGCTTAC 239
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 AAGAAGCTCTAAACACACTGTATCAAGACATATTGATGGT-----GATGGT 492
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 AAGGAAGCAACAAGCATCTTAGTCATGATGAGAATTAATGAAGCAAAATCATGATGGT 299
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY* 493 GAATGAAATATGTGTGATTTCTTTTAGCGGCTTAGGACAGATCTTCTCTA 552
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 GAATGCCGATACGTTGTGTGCTCGCTGATTAAGGGCTTGGAAACCGCACTACTCACTTT 359
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 553 GCCTCGGTTTTCTTTACCGGCTTTTAAACGATAGAGTCTTGTGTTGACCGAGGAA 612
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 GCTTCTGTGTTCTCTACGCTCTTCTGACGTATAGATCATCTTGTGTGACACCGCAAG 419
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 613 GACATGGATGATCTCTTTTTCGAGCCGTTTCTCGGTATGTCGTGGTTGCTACCTTTAGAT 672
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 GATATTGGTGTATCTCTATGCGAGCCATTTCCAGGTACTTCTGGTTGCTTCTCTCGAC 479
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 673 TTCCTATGACTGATCAGTTTGTATGATTAATCAAGATCATCTCGTTGTTATGGATAT 732
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 TTTCCATTGATGAATATGCTGATGATACCAAGGGATGCTCTGTTGTTACGAACA 539
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 733 ATGGTGAACAATCAGTGATGATGATGAGGAACTTTTCTCATCTTTATCTTCATCTT 792
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 ATGTTGGAAATCATTCCTCAACTCGATTTCTCCCGCCACATCTATATAGGCATAAC 599
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 793 GTTCATGATATTGGAGATCATGATAAGATGTTCTTCTGTGAAGAGGACCAACATCATC 852
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 CTTTCATGATTCAGGAGTAGTATAGATGTTCTTCTGCAAAAGATCAAGTTTGATT 659
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 853 GGGAAAGTCCCT 864
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 GACAAAGTCCCT 671
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
BF597909
LOCUS
DEFINITION
  BF597909 508 bp mRNA linear EST 06-DEC-2001
  Gm-c1056-656 5' similar to TR:081052 081052 T18E12.11 PROTEIN. [1]
  ;, mRNA sequence.
ACCESSION
  BF597909
VERSION
  BF597909.1
KEYWORDS
  GI:11690303
SOURCE
  soybean.
ORGANISM
  Glycine max
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
  1 (bases 1 to 508)
REFERENCE
  Shoemaker,R., Kelm,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
  ,A., Holla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
  Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
  ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
  ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
  ,R., Waterston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished (1999)
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available through: ResGen, Invitrogen Corp. 2130
  South Memorial Parkway Huntsville, AL 35801 For further information
  call: (800)-533-4363 or contact via email: ccufresgen.com
  Insert Length: 796 Std Error: 0.00
  High quality sequence stop: 411.
  Location/Qualifiers
    1..508
      /organism="Glycine max"
      /db_xref="taxon:3847"
      /clone="GENOME SYSTEMS CLONE ID: Gm-c1056-656"
      /clone_lib="Gm-c1056"
      /tissue_type="Whole seedling, 4 day old"
      /lab_host="DH10B"
      /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
        XhoI; The cDNA library was constructed from mRNA isolated
        from 4 day old seedling of P1468916. The seedlings were
        germinated in a growth chamber using germination paper.
        Complementary DNA was synthesized from mRNA using a primer
        consisting of a poly(dT) sequence with a XhoI restriction
        site. EcoRI adapters were ligated to the blunt-ended cDNA

```



fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 140 a 104 c 131 g 132 t 1 others  
ORIGIN  
Query Match 15.2%; Score 252.6; DB 12; Length 508;  
Best Local Similarity 72.3%; Pred. NO. 3.6e-63;  
Matches 327; Conservative 0; Mismatches 125; Indels 0; Gaps 0;  
Qy 1207 CCCAACACAAAGCGTCTGTCACATCTTTGACGCGGGTTACGCGGAGAACCTTAAG 1266  
Dbb 8 CCCAAGTCAAAAGCGGTACTGATGACATCCCTTAAACTCTGTTATTTGAAAAGGTGACA 67  
Qy 1367 AGTATGTAATCGGAATATCCGACATCAACTCGACAAATCATCCCTGTCATCAGCCGACC 1326  
Dbb 68 GACATGTATTTGGGAATATCTTACGGTGACGGGAGAGTGGTTGGCGTTTACACGCGGAGC 127  
Qy 1327 CAAGAAGGTTATCAGCAGACCGAAAAAAGATGCATATGCGAAAGCTCTTCGGGAAATG 1386  
Dbb 128 CACGAAGATATCAACAACACAGAGCAGATGCACACCAAAAGCTTGGCGAAGATG 187  
Qy 1387 TATCTTTTGTGTTGACAGATAATCTTGTGACAGTGTCTTGTCTACATTTTGGATATGTA 1446  
Dbb 188 TATCTTTTAAAGCTTAAGTGTGTGTGTACTACTCATGTCTACTTTTGGGTATG 247  
Qy 1447 GCTCAAGCTGTGAGCGGTTAAAGCTTTCATACATCTTATACACCCCAAAACCTTACACT 1506  
Dbb 248 GCTCANGGGCTTGGAGGTTTAAACCGTGGATTCTGTACAAACCTGAGAAATGGAAGCGC 307  
Qy 1507 CCGATCTCTGTGTGTCGCGGTATGTCGATGAGCGCTGTTCCTCACTCGCTCCATTC 1566  
Dbb 308 CCAGATCTCTCTGTCGACGTCGCAATGCAATGAGAGCGGTGTTCCATCTCTCTCCCTTC 367  
Qy 1567 TATGATTGTAAGCGAAACCGGTTATGACAGCGGAAACACTAGTTTCTCATGTGAGACAT 1626  
Dbb 368 TATGACTCAAGGCGACAGAGGAACACTCACAGCTGAACCTGTTCCTCATATGTAAGGCA 427  
Qy 1627 TGTGAGCAATACAGTCTGGGACTTTAAGCTAGT 1658  
Dbb 428 TGTGAAGATATGAGCTGGGGCTTAAAGCTTGT 459

RESULT 12  
BF009322  
Locus 561 bp mRNA linear EST 06-DEC-2001  
DEFINITION SS76g10.y1 Gm-cl064 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl064-91 5' similar to TR:081052 081052 T18E12.11 PROTEIN. [1]  
; mRNA sequence.

ACCESSION BF009322  
VERSION BF009322.1 GI:10709598  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 561)  
AUTHORS Shoemaker,R., Kaim,P., Vodkin,L., Erdelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,I., Underwood,K., Szeptoe,M., Theising,B., Allen,M., Bowers,V., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,B., Kohn,S., Shin,T., Jackson,V., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
TITLE Public Soybean EST Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewartson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccl@resgen.com  
High quality sequence stop: 422.

FEATURES  
source

1..561  
/organism="Glycine max"  
/cultivar="Williams"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl064-91"  
/clone\_lib="Gm-cl064"  
/tissue\_type="seedling epicotyls"  
/dev\_stage="2 week old"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from the epicotyls of 2 week old seedling for the cultivar Williams. The seedlings were germinated in a growth chamber, excised above the soil level, and the plants were placed in a 100 ppm solution of auxin for 24 hours prior to harvesting. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 137 a 112 c 109 g 202 t 1 others  
ORIGIN

Query Match 15.1%; Score 250.2; DB 12; Length 561;  
Best Local Similarity 67.0%; Pred. No. 1.9e-62;  
Matches 354; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

Qy 493 GAATCAAAATATGTTGTGCGATTTCTTTACCGCGCTTAGGAAACAGATACCTTTCTCA 552  
Dbb 34 CAGCTCAAAATATGTTGTGCGATTTCTTTACGCGCTTAGGAAATAGGATTTTGACCTA 93  
Qy 553 GCCTCGGTTTTCTTTACGCGCTTTTAAAGGATAGATCTTCTCTTGTGACCGAGGAA 612  
Dbb 94 GCTTCTGCTTTCTCTATGCTCTCTCTCACTAACCGGTCTTCTACTGCTTCACTGAGCG 153  
Qy 613 CACATCGATCATCTCTTTTTCGAGCGCTTTCTCGGTATGTCGGTGTGCTTACCTTTAGAT 672  
Dbb 154 GATATGGTTGATCTCTTTTGTGAGCCATTCAGATCTCTCTGGCTTCTCTAGTAGAC 213  
Qy 673 TTCCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732  
Dbb 214 TTTTCCCTTAAATGCTCAGTTTAAATTTTTCAGTCAGATTTCTGATCATTTGTTATGGAAA 273  
Qy 733 ATGGTGAAGATCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 792  
Dbb 274 ATGCTGAAAGTAAAGACACATGATGATGATGATGATGATGATGATGATGATGATGAT 333  
Qy 793 CTTTATGATTTATGAGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 852  
Dbb 334 GCCCATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 393  
Qy 853 GGGAAAGTCCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 912  
Dbb 394 CAGATCGTCCCATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 453  
Qy 913 ATACCGGGTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972  
Dbb 454 ATGCCATCTTTTGAGCAGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 513



[illegible]

Db	156	TTTGTCCCTTCTGTATTCTTAATGCCATCTTTTGACGAGAACTGAATGATCTCTTTCCCA	215
Qy	952	CAGAAAGGACTGCTTTTCATCACITTAGTGTAGTATCTTTTACCACCACTAACCAAGTA	101
Db	216	AATAAGGAACACGTGTCCATTTTGTGAGTAGGTATCTGTGCACCCCTACAACACAGTGTG	275

Qy	1012	TGGGGCTTACTCAGTACGAACTTATTCGCATCGCATGAGAAGATTGGG	107
Db	276	TGGGGACTTGCTCAGATACTATCAAGCTTATTATCTAAAGCTGATGAAGAGTGGCC	335

QY	1072	ATTCAAGTAAGAGTTTTTCGATGAAGACCGGGTCCATTTTCAGCATCTGATGGATCAGATT	113
Db	336	ATACAATTAGAGTGTTGACACTGAACCTGGTCCGTTTCAACATGTATTGGATCAGATC	395
QV	1132	TGATCTTGCTACTCAAAGAAGAAAGTTCTAGCTAGTAGTACACACACTAGTGCAGACAGCTC	119

1132	1CATCTGTGTA	CTCAAAAGAGAAAC	TCTACCTGGAAGT	TAGACACACTAGT	GAGAGATCT	1133
QY						
Db						
396	TTAGCTGTGTA	CTTTGAAGAAGAATA	TTTTGCGTGATGT	TAAACCCACGAGAT	GCTCTACC	455
QY						
1192	CGCATGTTTA	TACCCCCCAACACA	AAAGCGGTGCT	TGTGCACACTCTTT	GAACCGCGGTAC	1253
QY						
456	AATTTCATC	AGGAATTC	CCCAAGTCAA	AAAGCGGTACT	GATGACACTCTT	515
Db						

[illegible]

RESULT 14	AV831440	LOCUS	AV831440	314 bp	mRNA	linear	EST 01-APR
DEFINITION	AV831440	RAFL9 Arabidopsis thaliana cDNA clone RAFL09-89-D07 5'					

DEFINITION	Arabidopsis thaliana cDNA clone RAFL09-89-DU/1
ACCESSION	AF831440
VERSION	AF831440.1
KEYWORDS	EST.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 314) Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, and Shinozaki, K.
TITLE	Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL	Unpublished (2002)
COMMENT	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060

[illegible]

clone is in a modified p Bluescript vector. Please visit our website ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further information.

**FEATURES**

Location/Qualifiers

details.

```
source
1. .314
/organism="Arabidopsis thaliana"
```

```
/db_xref="taxon:3702"
/clone_lib="RAFL09-89-D07"
/clone_lib="RAFL9"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/notice="Site_1: BamHI; Site_2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
BASE COUNT      88 a      64 g      86 t      3 others
ORIGIN
Query Match      14.2%; Score 235.6; DB 10; Length 314;
Best Local Similarity 96.5%; Pred. NO. 3.1e-58;
Matches 250; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
QY 1 ATGGATCAGATTCTACAGGAGAGATCGTCTCCGATCAGAACCACT--ACCGGGGGTT 58
DB 56 ATGGATCAGATTCTACAGGAGAGATCGTCTCCGATCAGAACCACTTACCGGGGGTT 115
QY 59 CAAGTCGGTTAAATTTCTCCGAATCTTCAAAATGAAGTATCTCAGTCCGGTACGATCA 118
DB 116 CAAGTCGGTTAAATTTCTCCGAATCTTCAAAATGAAGTATCTCAGTCCGGTACGATCA 175
QY 119 AGCTCAGAGAACCTTCACTACTTCTGATAGTCTCTCTCTAGTAGTAGTCTCAA 178
DB 176 AGCTCAGAGAACCTTCACTACTTCTGATAGTCTCTCTCTAGTAGTAGTCTCAA 235
QY 179 TGATCTTTTCAACACCACTTGATTTCAATCGGATATGGTTTCGCGGAGCTAGAG 238
DB 236 TGATCTTTTCAACACCACTTGATTTCAATCGGATATGGTTTCGCGGAGCTAGAG 295
QY 239 TTCTCGACCGCGAGTTT 257
DB 296 TTCTCGACCGCGAGTTT 314

RESULT 15
B18679/c
Locus      B18679      786 bp      DNA      linear      CSS 09-AUG-1997
DEFINITION F16A14-T7 IGF Arabidopsis thaliana genomic clone F16A14, DNA
sequence.
ACCESSION  B18679
VERSION    B18679.1 GI:2316583
KEYWORDS   CSS
SOURCE     thale cress.
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases) 1 to 786)
Feng,J., Dewar,K., Buehler,E., Kim,C., Li,V., Shinn,P., Sun,H. and
Ecker,J.
BAC End Sequences at ATGC
Unpublished (1997)
Other GSSs: F16A14-sp6
Contact: Ecker J
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 64
High quality sequence stop: 150.
Location/Qualifiers
1. .786
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
```

```
/clone="F16A14"
/clone_lib="IGF"
/sex="hermaphrodite"
/notes="Vector: BelobAC11; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"
BASE COUNT      189 a      204 c      142 g      232 t      19 others
ORIGIN
Query Match      14.0%; Score 232.4; DB 17; Length 786;
Best Local Similarity 63.4%; Pred. NO. 3.8e-57;
Matches 371; Conservative 0; Mismatches 210; Indels 4; Gaps 2;
QY 1054 GCGGATCAGAACGATTCGGATTAAGTATTCGATGAAGCCGGGTTCGATTTTCA 1113
DB 715 GCGGAGGAGGNGGTTATTAGAGAAGGGGTAGTTTCAGCATTCGAGGGGAAATNTAA 656
QY 1114 CATGTGATGGATCAGATTTTCATCTTGTACTCAAAACAGAAACCTTCACTGAGTAC 1173
DB 655 CAGGNAAGGATCAAGTTGTAACTTGGACAGGAAGAGAG--ACCTTTTGCCTGAACAAG 597
QY 1174 ACACATAGTGGAGAGATCTCGC---CATGTTAATACCCCAACACAAAGCGCTTGTG 1230
DB 596 GCACAAGAGAGCCAAAGTCATATATCAAGAAGCCAGACACTTANAGCTGTTCTNAG 537
QY 1231 ACATCTTTTCAACCGCGGTTACGCGGAGAACTTAAAGAGTATGTTGGGAATATCCGACA 1290
DB 536 ACATCTTNGTATCCAGAGTACTCTGANACCTTANAGAAGCATGTTTGGGAACGCGAGT 477
QY 1291 TCAACTGGAGAAATCATGCGTTCATCAGCCAGCCAGCAAGGTTATCAGCAGACCGAA 1350
DB 476 TCAGCAGGGGAGATCATTTGAAGTTTATCAGCAAGTGGAGAAAGGTTTCAACANACAG 417
QY 1351 AAAAAATGCATAATGGCAAGCTCTTGGGAAATGTATCTTTTGACTTTTCACAGATAAT 1410
DB 416 AAGAANTACACCAACCAAGGCGCTCGCGGAGANGTATCTTCTAAGTCTAAGTGTAA 357
QY 1411 CTGTGACAAAGTCTTGGTCTACATTTGGATATGATGATCAAGGTCCTTGGAGGTTAAAG 1470
DB 356 ATAGTCACAAGTCAAGGCTACATTTGGATATGTCCTCATATTTCTTCCAGCATTAAG 297
QY 1471 CATTGATATCTATATACCCGAAACCGTACAACTCCCGATCTTGGTGTGGTTCGGGCT 1530
DB 296 CCATGGTTACTTTATCAGCCAGCAGGTCTGATACAGTCTCTGATCCACAGNGTATTCAATCC 237
QY 1531 ATGTCGATGGAGCTTGTTCCTCACTCGCTCCATTCATGATTCGTAAGCGAAACCGGT 1590
DB 236 AGCTGATGGAGCATATGTCACCTTACTTTCCTCTCATGGATGTGAACCCGAATGGGA 177
QY 1591 ATTGACACGGGAACTAGTTTCTCTCATGTGAGAGACATTGTGAGGAT 1635
DB 176 ACTAAGTCCGGGAAGTAGTTCCTCTGNGNAAGCATTTGTCAGCAT 132
```

Search completed: March 5, 2003, 12:33:02  
Job time : 2208 secs